

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: December 22, 2004, 06:23:58 ; Search time 191 Seconds

(without alignments)  
93.385 Million cell updates/sec

Title: US-09-732-348B-34

Perfect score: 54

Sequence: 1 XXCXXXXXXXXXXXXXXXXXXXXXXXXX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	35.2	103	1	YA40 MYCPN
2	18	33.3	39	2	Q7PER7 anopheles g
3	18	33.3	49	2	Q7UWS9 rhodopirell
4	18	33.3	58	2	Q7PE97 anopheles g
5	18	33.3	61	2	Q86YX3 homo sapien
6	18	33.3	79	1	CMC1 EIMAC
7	18	33.3	87	2	Q7PE39 anopheles g
8	18	33.3	88	2	Q94A28 arabidopsis g
9	18	33.3	111	2	Q6K499 ratius norv
10	18	33.3	113	2	Q6K499 ratius norv
11	18	33.3	113	2	BAD22270
12	18	33.3	116	2	Q40635
13	18	33.3	121	1	CDS9_SHV21
14	18	33.3	124	2	Q9GM28 macaca fasc
15	18	33.3	129	2	Q6LIG2 dirosophila
16	18	33.3	131	2	Q8N723 homo sapien
17	18	33.3	135	2	Q63317 ratius norv
18	18	33.3	137	2	Q24987 giardia lam
19	18	33.3	138	2	Q8C2X1 mus musculu
20	18	33.3	139	2	Q6LBB8 oligotropha
21	18	33.3	139	2	CAG28441
22	18	33.3	141	2	Q96AC2 homo sapien
23	18	33.3	141	2	Q8BLC3 mus musculu
24	18	33.3	141	2	Q9J396 mus musculu
25	18	33.3	141	2	AAB58599 mus muscu
26	18	33.3	141	2	AAB58599 mus muscu
27	18	33.3	157	1	UL42_HCVNA
28	18	33.3	157	2	Q6ZP52 homo sapien
29	18	33.3	157	2	BAC85273
30	18	33.3	160	2	Q7OF28 anopheles g
31	18	33.3	162	2	Q88EY6 pseudomonas

32	18	33.3	164	2	Q6ZWT4	Q6ZWT4 homo sapien
33	18	33.3	164	2	BAC85518	BAC85518 homo sapi
34	18	33.3	165	2	Q8N264	Q8N264 homo sapien
35	18	33.3	167	2	Q24969	Q24969 giardia lam
36	18	33.3	167	2	Q6IEP8	Q6IEP8 cryza sativ
37	18	33.3	170	2	Q9BIL7	Q9BIL7 giardia lam
38	18	33.3	173	2	Q77337	Q77337 equine arte
39	18	33.3	173	2	Q9WAI1	Q9WAI1 equine arte
40	18	33.3	173	2	Q9WD19	Q9WD19 equine arte
41	18	33.3	173	2	Q9WD20	Q9WD20 equine arte
42	18	33.3	173	2	Q9WD26	Q9WD26 equine arte
43	18	33.3	173	2	Q9WD27	Q9WD27 equine arte
44	18	33.3	173	2	Q9WD29	Q9WD29 equine arte
45	18	33.3	173	2	Q9WD39	Q9WD39 equine arte

## ALIGNMENTS

```

RESULT 1
YA40 MYCPN          STANDARD;          PRT;          103 AA.
ID  YA40 MYCPN
AC  P75074;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DE  05-JUL-2004 (Rel. 44, Last annotation update)
DE  Hypothetical protein MPN040 (B01_orf103b) .
GN  Hypothetical protein MPN040; ORFNames=MP114;
OS  Mycoplasma pneumoniae.
OC  Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX  NCBI_TaxID=2104;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 29342 / M129;
RX  MEDLINE=97105885; Pubmed=8948633;
RA  Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
   Hermann R.;
RT  "Complete sequence analysis of the genome of the bacterium Mycoplasma
   pneumoniae."
RL  Nucleic Acids Res. 24:4420-4449(1996).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.1sb-sib.ch/announce/
   or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; AB000014; AAB95762.1; -.
DR  PIR; S73440; S73440.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE 103 AA; 11291 MW; 62E96847339936 CRC64;

Query Match          35.2%; Score 19; DB 1; Length 103;
Best Local Similarity 13.6%; Pred. No. 6.7e+02;
Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY  3 CXXXXXXXXXXXXXXXXXXXXX 24
DB  53 CSSISFCSLASSARLRYSST 74

RESULT 2
Q7PER7              PRELIMINARY;          PRT;          39 AA.
ID  Q7PER7;
AC  Q7PER7;
DT  01-MAR-2004 (TEMBLrel. 26, Created)
DT  01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT  01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE  ENSANGP0000023374.
GN  Name=ENSANGG0000020343;
OS  Anopheles gambiae str. PEST.

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: AAA801008326; EAA45619.1; -;  
 SQ SEQUENCE 39 AA; 4317 MW; 6963C9D228D7C8C7 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 39;  
 Best Local Similarity 28.6%; Pred. No. 6.9e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 DB 2 CAATASC 8

## RESULT 3

O7UMS9 PRELIMINARY; PRT; 49 AA.  
 AC O7UMS9;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=RA1825;  
 OS Rhodopirellula baltica  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Glocker F.O., Kube W., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,  
 RA Schleuter H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 KM EMBL: BX294135; CAD72283.1; -;  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 49 AA; 5561 MW; D47FD5A9BFF7BA40 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 49;  
 Best Local Similarity 28.6%; Pred. No. 8e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 DB 3 CSSSASC 9

## RESULT 4

O7PE97 PRELIMINARY; PRT; 58 AA.  
 AC O7PE97;  
 DT 01-MAR-2004 (TREMBLrel. 26, Created)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE ENSANGP0000022719 (Fragment).  
 GN Name=ENSANGG0000021658;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: AAA801004888; EAA45826.1; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 58 58  
 SQ SEQUENCE 58 AA; 6678 MW; BDD2E5D73967DAB5 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 58;  
 Best Local Similarity 28.6%; Pred. No. 9e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 DB 28 CAATASC 34

## RESULT 5

O86YX3 PRELIMINARY; PRT; 61 AA.  
 AC O86YX3;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE MTE.  
 GN Name=MTE;  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yu L.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Metallothioneins have a high content of cysteine  
 CC residues that bind various heavy metals (By similarity).  
 CC -!- SIMILARITY: Belongs to the metallothionein superfamily. Family 1.  
 DR EMBL: AF348997; AAO32957.1; -;  
 DR HSSP: P02795; 1MHU  
 DR GO: GO:0046872; Fimetal ion binding; IEA.  
 DR InterPro: IPR003019; Metallothion\_1.  
 DR InterPro: IPR000006; Metallothion\_1.  
 DR Pfam: PF00131; Metallothio\_1.  
 DR PRINTS: PR00860; MTEVERTBRARE.  
 KW Metal-binding; Metal-thiolate cluster.  
 SQ SEQUENCE 61 AA; 6165 MW; 7F71CE7D1C37762A CRC64;

Query Match 33.3%; Score 18; DB 2; Length 61;  
 Best Local Similarity 28.6%; Pred. No. 9.3e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 DB 7 CSTSSSC 13

## RESULT 6

CMC1\_EIMAC STANDARD; PRT; 79 AA.  
 AC P21959;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE EAMZP30-47 protein (Fragment).  
 GN Name=CMC17;  
 OS Eimeria acervulina.  
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;  
 OC Bimeria.  
 OX NCBI\_TaxID=5801;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20184343; PubMed=1690144;  
 RA Jenkins M.C., Lillehoj H.S., Barta J.R., Danforth H.D.,  
 RA Stroblein D.A.;  
 RT "Primeria acervulina: cloning of a cDNA encoding an immunogenic region  
 RT of several related merizote surface and rhotopy proteins";  
 RL Exp. Parasitol. 70:353-362(1990).  
 CC -1- SUBCELLULAR LOCATION: Surface membrane and internal rhotopies.  
 CC -1- DEVELOPMENTAL STAGE: Merizote.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL, M37843; AAA62796.1; -.  
 KW Antigen; Membrane; Merizote.  
 FT NON TER 1  
 FT NON TER 79  
 SQ SEQUENCE 79 AA; 8703 MW; 0B2B6CDE65FB4330 CRC64;  
 Query Match 33.3%; Score 18; DB 1; Length 79;  
 Best Local Similarity 28.6%; Pred. No. 1.1e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Oy 3 CXXXXXXC 9  
 Db 30 CXXXXXXC 36  
 RESULT 7  
 Q7PE39 PRELIMINARY; PRT; 87 AA.  
 AC Q7PE39;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DE EMBL=2004 (TREMBlrel. 26, Last annotation update)  
 GN EMBL=2004 (TREMBlrel. 26, Last annotation update)  
 OS Name=EMBL=2004 (TREMBlrel. 26, Last annotation update)  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL, AAB01002921; EAA45886.1; -.  
 FT NON TER 1  
 FT NON TER 87  
 SQ SEQUENCE 87 AA; 10053 MW; 45F57BD9A1AE0FC3 CRC64;  
 Query Match 33.3%; Score 18; DB 2; Length 87;  
 Best Local Similarity 28.6%; Pred. No. 1.2e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Oy 3 CXXXXXXC 9  
 Db 28 CAATASC 34  
 RESULT 8  
 Q94A28 PRELIMINARY; PRT; 88 AA.  
 AC Q94A28;

DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE A7492230/T1014 60 (Hypothetical protein).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,  
 RA Banth J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,  
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,  
 RA Kamaya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,  
 RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,  
 RA Davis R.W., Theologis A., Ecker J.R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22088475; PubMed=12093376;  
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation";  
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.,  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banth J.,  
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamaya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AY045581; AAK7339.1; -.  
 DR EMBL, AY084624; AAM61187.1; -.  
 DR EMBL, AY094030; AAM61186.1; -.  
 DR GO: GO:0006952; P: defense response; IEA.  
 DR InterPro: IPR003614; Kof1.  
 DR SMART: SM00505; Kof1.1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 88 AA; 9071 MW; 2345DAB2B6E96A8 CRC64;  
 Query Match 33.3%; Score 18; DB 2; Length 88;  
 Best Local Similarity 28.6%; Pred. No. 1.2e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Oy 3 CXXXXXXC 9  
 Db 40 CTSSSTC 46  
 RESULT 9  
 Q63316 PRELIMINARY; PRT; 111 AA.  
 AC Q63316;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Ly6-A antigen (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RA MEDLINE=90152758; PubMed=2154400;
RT Friedman S., Palfrey R.G.E., Sirlin S., Haemmerling U.;
RT "Analysis of three distinct ly6-A-related cDNA sequences isolated from
RT rat kidney.";
RL Immunogenetics 31:104-111(1990).
DR EMBL; M30692; AAA1545.1; -.
DR InterPro; IPR001526; LY6_UPAR.
DR Pfam; PF00021; UPAR_LY6; 1.
DR SMART; SM00134; LY; 1.
FT NON TER
SQ SEQUENCE 111 AA; 11755 MW; B0A42D1B3ED148C4 CRC64;

Query Match
Best Local Similarity 33.3%; Score 18; DB 2; Length 111;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 17 CSASAC 23

RESULT 10
06K499 PRELIMINARY; PRT; 113 AA.
AC 06K499;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein OJ1595_D08.20.
GN Name=OJ1595_D08.20;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005574; BAD22270.1; -.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 12326 MW; C2A6C98129828244 CRC64;

Query Match
Best Local Similarity 33.3%; Score 18; DB 2; Length 113;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 35 CSASAC 41

RESULT 11
BAD22270 PRELIMINARY; PRT; 113 AA.
ID BAD22270;
DT 01-JUN-2004 (TReMBLrel. 27, Created)
DT 01-JUN-2004 (TReMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein OJ1595_D08.20.
GN OJ1595_D08.20.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;

```

```

RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
RT clone:OJ1595_D08.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005574; BAD22270.1; -.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 12326 MW; C2A6C98129828244 CRC64;

Query Match
Best Local Similarity 33.3%; Score 18; DB 2; Length 113;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 35 CSASAC 41

RESULT 12
040635 PRELIMINARY; PRT; 116 AA.
ID 040635;
AC 040635;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE CD59 protein (viral CD59 antigen).
GN Name=prf15; Synonyms=VCD59;
OS Saimiriine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-488;
RA MEDLINE=98037620; PubMed=9371569;
RX Knäppe A., Hiller C., Thurnau M., Wiltmann S., Hofmann H.,
RA Fleckenschein B., Fickenscher H.;
RT "The superantigen-homologous viral immediate-early gene ie14/vsa9 in
RT herpesvirus saimiri-transformed human T cells.";
RL J. Virol. 71:9124-9133(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C488;
RX MEDLINE=22918177; PubMed=14554077;
RA Essner A., Thurnau M., Wiltmann S., Fickenscher H.;
RT "The genome of herpesvirus saimiri C488 which is capable of
RT transforming human T cells.";
RL Virology 314:471-487(2003).
DR EMBL; Y13183; CAA73629.1; -.
DR EMBL; AJ410493; CAC84310.1; -.
DR HSP; P13987; 1ERG.
DR InterPro; IPR003632; LY-6_CD59.
DR InterPro; IPR001526; LY6_UPAR.
DR Pfam; PF00021; UPAR_LY6; 1.
DR ProDom; PD003128; LY-6_CD59; 1.
DR SMART; SM00134; LY; 1.
DR PROSITE; PS00983; LY6_UPAR; 1.
SQ SEQUENCE 116 AA; 13172 MW; 68F35CE923FF189 CRC64;

Query Match
Best Local Similarity 33.3%; Score 18; DB 2; Length 116;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 32 CTTSTSC 38

RESULT 13
CD59_SHV21 STANDARD; PRT; 121 AA.
ID CD59_SHV21;
AC 000996;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)

```

```

DT 05-JUL-2004 (Rel. 44, last annotation update)
DE Surface glycoprotein CD59 homolog precursor.
GN Name=15;
OS Saimiriine herpesvirus 2 (strain 11) (SaHV-2) (Herpesvirus saimiri).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333686; Pubmed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wiltmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
RN [2]
RP SIMILARITY TO CD59.
RX MEDLINE=92410640; Pubmed=1382344;
RA Albrecht J.-C., Nicholas J., Cameron K.R., Newman C., Fleckenstein B.,
RA Honess R.W.;
RT "Herpesvirus saimiri has a gene specifying a homologue of the cellular
RT membrane glycoprotein CD59.";
RL Virology 190:527-530(1992).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (by
CC similarity).
CC -1- SIMILARITY: Contains 1 UPAR/Ly6 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: X64346; CAA45638.1; -.
DR EMBL: X64273; CAA45565.1; -.
DR HSP: P13987; 1ERG.
DR InterPro: IPR003632; Ly-6-CD59.
DR InterPro: IPR001526; Ly6_UPAR.
DR Pfam: PF00021; UPAR_Ly6; 1.
DR ProDom: PD003128; Ly-6_CD59; 1.
DR SMART: SM00134; LU; 1.
DR PROSITE: PS00983; LY6_UPAR; 1.
KW Glycoprotein; GPI-anchor; Lipoprotein; Signal.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 96 Surface glycoprotein CD59 homolog.
FT .PROPEP 97 121 Removed in mature form (Potential).
FT DOMAIN 20 104 UPAR/Ly6.
FT DISULFID 22 45 By similarity.
FT DISULFID 25 32 By similarity.
FT DISULFID 38 58 By similarity.
FT .DISULFID 64 82 By similarity.
FT .DISULFID 83 88 By similarity.
FT LIPID 96 96 GPI-anchor amidated asparagine (by host
FT (Potential)).
FT CARBOHYD 24 24 N-linked (GlcNAc...)(Potential).
SQ SEQUENCE 121 AA; 13814 MW; 56A8BDC07B953AA3 CRC64;
Query Match 33.3%; Score 18; DB 1; Length 121;
Best Local Similarity 28.6%; Pred. No. 1.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
Db 32 CTTSTSC 38
RESULT 14
O9GM28 PRELIMINARY; PRT; 124 AA.
AC O9GM28;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)

```

```

DT 01-MAR-2001 (T-EMBLrel. 16, last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain parietal lobe;
RX MEDLINE=21458551; Pubmed=11574149;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirata M., Suto Y.,
RA Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K.;
RT "Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
RT chromosomes.";
RL Gene 275:31-37(2001).
DR EMBL: AB049858; BAB16744.1; -.
KW Hypothetical protein.
SQ SEQUENCE 124 AA; 13904 MW; B87262632AF6D149 CRC64;
Query Match 33.3%; Score 18; DB 2; Length 124;
Best Local Similarity 28.6%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
Db 16 CAASSTC 22
RESULT 15
O6ILG2 PRELIMINARY; PRT; 129 AA.
ID O6ILG2;
AC O6ILG2;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, last annotation update)
DE H0C09485.
GN ORFNames=H0C09485;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX Pubmed=14709175;
RA Hild M., Beckmann B., Haas S., Koch B., Solovjev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hohnleisel J., Pato R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:R3-R3(2003).
CC -1- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (Tpa) entry.
DR EMBL: BK02054; DAA02899.1; -.
SQ SEQUENCE 129 AA; 13635 MW; 1E5850AE37A3F647 CRC64;
Query Match 33.3%; Score 18; DB 2; Length 129;
Best Local Similarity 28.6%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
Db 91 CSSTSSC 97
Search completed: December 22, 2004, 06:33:53
Job time : 195 secs

```

**This Page Blank (us,**

**This Page Blank (uspto)**



```

XX
SQ Sequence 103 AA;
Query Match 35.2%; Score 19; DB 7; Length 103;
Best Local Similarity 13.6%; Pred. No. 1.2e+04;
Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 3 CXXXXXXCXXXXXXXXXXXXX 24
Db 53 CSSISFCSLASSSARLRYSSTH 74

RESULT 2
AAW65945
ID AAW65945 standard; peptide; 10 AA.
XX
AC AAW65945;
XX
DT 12-NOV-1998 (first entry)
XX
DE .Molecule V.
XX
KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;
KM BDNF; NT-3; conformation; promoter.
XX
OS Synthetic.
OS Mammalia.
XX
PN CA205045-A.
XX
PD 12-MAY-1998.
XX
PF 09-MAY-1997; 97CA-02205045.
XX
PR 12-NOV-1996; 96CA-02190296.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;
XX
DR WPI; 1998-457650/40.
XX
PT Inhibition of neurotrophin activity - using factor that interferes with
PT sub-unit interaction.
XX
PS Disclosure; Page 13; 103pp; English.
XX
CC The invention relates to a method for reducing the biological activity of
CC a multimeric protein having at least 2 promoters. It comprises (a)
CC providing a factor that interacts with at least one portion of at least
CC one of the promoters which associates with a portion of the other
CC promoter in the absence of the factor; and (b) mixing the factor with the
CC multimeric protein so that the factor interacts with the portion(s) and
CC disrupts association of at least a portion of the promoters. The method
CC and compounds are useful for inhibiting neurotrophin-mediated activities
CC selected from neurotrophin receptor binding, neuron survival, neurite
CC outgrowth and epileptic effects. The present sequence is shown in the
CC specification
XX
SQ Sequence 10 AA;
Query Match 33.3%; Score 18; DB 2; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
Db 1 CAAAAAC 7

RESULT 3
AAW6291
ID AAW6291 standard; peptide; 10 AA.

```

```

XX
AC AAW6291;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 2485.
XX
KM Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB004776.
XX
PR 13-DEC-1999; 99GB-00029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX
PS Example 4; Page 397; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX
SQ Sequence 10 AA;
Query Match 33.3%; Score 18; DB 4; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 4
AAW6207
ID AAW6207 standard; peptide; 10 AA.
XX
AC AAW6207;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 2401.
XX
KM Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB004776.
XX
PR 13-DEC-1999; 99GB-00029464.
XX
PA (PROT-) PROTEOM LTD.

```



```
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX
PS Example 4; Page 385; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX
SQ Sequence 10 AA;

Query Match          33.3%; Score 18; DB 4; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 2 CSSSTTC 8

RESULT 5
ID AAG96231 standard; peptide; 10 AA.
XX
AC AAG96231;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 2425.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GH004776.
XX
PR 13-DEC-1999; 99GB-00029464.
XX
PA - (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX
PS Example 4; Page 389; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX
SQ Sequence 10 AA;
```

```
Query Match          33.3%; Score 18; DB 4; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 1 CSSSTTC 7

RESULT 6
ID AAG96205 standard; peptide; 10 AA.
XX
AC AAG96205;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 2399.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GH004776.
XX
PR 13-DEC-1999; 99GB-00029464.
XX
PA - (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX
PS Example 4; Page 385; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX
SQ Sequence 10 AA;

Query Match          33.3%; Score 18; DB 4; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 4 CSSSTTC 10

RESULT 7
ID AAG96349 standard; peptide; 10 AA.
XX
AC AAG96349;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 2543.
```

```

XX KM Human; complementary peptide; ligand; drug discovery; drug design.
XX OS Homo sapiens.
XX PN WO200142277-A2.
XX PD 14-JUN-2001.
XX PF 13-DEC-2000; 2000WO-GB004776.
XX PR 13-DEC-1999; 99GB-00029464.
XX PA (PROT-) PROTEOM LTD.
XX PI Roberts GW, Heal JR;
XX DR WPI; 2001-408419/43.
XX PT A set of peptide ligands consisting of specific complementary peptides to
XX PT proteins encoded by genes of the human genome, useful in an assay for
XX PT screening and identifying of one or more novel peptides which are drug
XX PT candidates or pro-drugs.
XX PS Example 4; Page 406; 646pp; English.
XX CC The invention relates to a set of complementary peptide ligands generated
XX CC from the human genome. The complementary peptides interact with their
XX CC relevant target proteins encoded in the human genome. They can be used as
XX CC reagents in drug discovery and as lead ligands to facilitate drug design
XX CC and development. The present sequence is a complementary peptide provided
XX CC in the specification
XX SQ Sequence 10 AA:

Query Match 33.3%; Score 18; DB 4; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0
QY 3 CXXXXXC 9
DB 4 CSSSTTC 10

RESULT 8
AAG96191
ID AAG96191 standard; peptide; 10 AA.
XX AC AAG96191;
XX DT 18-SEP-2001 (first entry)
XX DE Human complementary peptide; SEQ ID NO: 2385.
XX KW Human; complementary peptide; ligand; drug discovery; drug design.
XX OS Homo sapiens.
XX PN WO200142277-A2.
XX PD 14-JUN-2001.
XX PF 13-DEC-2000; 2000WO-GB004776.
XX PR 13-DEC-1999; 99GB-00029464.
XX PA (PROT-) PROTEOM LTD.
XX PI Roberts GW, Heal JR;
XX DR WPI; 2001-408419/43.
XX PT A set of peptide ligands consisting of specific complementary peptides to
XX PT proteins encoded by genes of the human genome, useful in an assay for
XX PT proteins encoded by genes of the human genome, useful in an assay for

```

```

PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX
PS Example 4; Page 383; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX
SQ Sequence 10 AA;

Query Match          33.3%; Score 18; DB 4; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
   |
Db 3 CSSSTTC 9

RESULT 9
AAG6339
ID AAG6339 standard; peptide; 10 AA.
XX
AC AAG6339;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, ligand; drug discovery; drug design.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB004776.
XX
PR 13-DEC-1999; 99GB-00029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberte GW, Heal JR;
XX
DR WI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome; useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX
XX Example 4; Page 404; 646pp; English.
XX
PS
XX
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX
SQ Sequence 10 AA;

Query Match          33.3%; Score 18; DB 4; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
   |
Db 3 CSSSTTC 9

```

DB 2 CSSSTTC 8

RESULT 10  
AAG96363  
ID AAG96363 standard; peptide; 10 AA.  
XX AAG96363;  
AC AAG96363;  
XX 18-SEP-2001 (first entry)  
XX Human complementary peptide; SEQ ID NO: 2557.  
XX Human; complementary peptide; ligand; drug discovery; drug design.  
XX Homo sapiens.  
XX WO200142277-A2.  
XX 14-JUN-2001.  
XX 13-DEC-2000; 2000WO-GB004776.  
XX 13-DEC-1999; 99GB-00029464.  
XX (PROT-) PROTEOM LTD.  
XX Roberts GW, Heal JR;  
XX WPI; 2001-408419/43.  
XX A set of peptide ligands consisting of specific complementary peptides to  
PT proteins encoded by genes of the human genome, useful in an assay for  
PT screening and identifying of one or more novel peptides which are drug  
PT candidates or pro-drugs.  
XX Example 4; Page 408; 646pp; English.  
XX The invention relates to a set of complementary peptide ligands generated  
CC from the human genome. The complementary peptides interact with their  
CC relevant target proteins encoded in the human genome. They can be used as  
CC reagents in drug discovery and as lead ligands to facilitate drug design  
CC and development. The present sequence is a complementary peptide provided  
CC in the specification  
XX Sequence 10 AA;  
SQ

Query Match 33.3%; Score 18; DB 4; Length 10;  
Best Local Similarity 28.6%; Pred. No. 5.8e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 3 CXXXXXC 9  
DB 2 CSSSTTC 8

RESULT 11  
AAG96333  
ID AAG96333 standard; peptide; 10 AA.  
XX AAG96333;  
AC AAG96333;  
XX 18-SEP-2001 (first entry)  
XX Human complementary peptide; SEQ ID NO: 2527.  
XX Human; complementary peptide; ligand; drug discovery; drug design.  
XX Homo sapiens.  
XX WO200142277-A2.  
XX 14-JUN-2001.  
PD

XX 13-DEC-2000; 2000WO-GB004776.  
XX 13-DEC-1999; 99GB-00029464.  
XX (PROT-) PROTEOM LTD.  
XX Roberts GW, Heal JR;  
XX WPI; 2001-408419/43.  
XX A set of peptide ligands consisting of specific complementary peptides to  
PT proteins encoded by genes of the human genome, useful in an assay for  
PT screening and identifying of one or more novel peptides which are drug  
PT candidates or pro-drugs.  
XX Example 4; Page 403; 646pp; English.  
XX The invention relates to a set of complementary peptide ligands generated  
CC from the human genome. The complementary peptides interact with their  
CC relevant target proteins encoded in the human genome. They can be used as  
CC reagents in drug discovery and as lead ligands to facilitate drug design  
CC and development. The present sequence is a complementary peptide provided  
CC in the specification  
XX Sequence 10 AA;  
SQ

Query Match 33.3%; Score 18; DB 4; Length 10;  
Best Local Similarity 28.6%; Pred. No. 5.8e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 3 CXXXXXC 9  
DB 2 CSSSTTC 8

RESULT 12  
AAG96193  
ID AAG96193 standard; peptide; 10 AA.  
XX AAG96193;  
AC AAG96193;  
XX 18-SEP-2001 (first entry)  
XX Human complementary peptide; SEQ ID NO: 2387.  
XX Human; complementary peptide; ligand; drug discovery; drug design.  
XX Homo sapiens.  
XX WO200142277-A2.  
XX 14-JUN-2001.  
XX 13-DEC-2000; 2000WO-GB004776.  
XX 13-DEC-1999; 99GB-00029464.  
XX (PROT-) PROTEOM LTD.  
XX Roberts GW, Heal JR;  
XX WPI; 2001-408419/43.  
XX A set of peptide ligands consisting of specific complementary peptides to  
PT proteins encoded by genes of the human genome, useful in an assay for  
PT screening and identifying of one or more novel peptides which are drug  
PT candidates or pro-drugs.  
XX Example 4; Page 384; 646pp; English.  
XX The invention relates to a set of complementary peptide ligands generated  
CC from the human genome. The complementary peptides interact with their

CC relevant target proteins encoded in the human genome. They can be used as  
 CC reagents in drug discovery and as lead ligands to facilitate drug design  
 CC and development. The present sequence is a complementary peptide provided  
 CC in the specification  
 XX  
 SQ Sequence 10 AA;

Query Match 33.3%; Score 18; DB 4; Length 10;  
 Best Local Similarity 28.6%; Pred. No. 5.8e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 DB 4 CSSSTTC 10

RESULT 13  
 AAW54432  
 ID AAW54432 standard; peptide; 12 AA.  
 XX  
 AC -AAW54432;

DT 15-SEP-1998 (first entry)

XX Human PS112 protein synthetic peptide #7.

XX Prostate; disease; PS112 gene; detection; diagnosis; cancer; treatment;  
 KW antibody.

XX Homo sapiens.  
 OS Synthetic.

XX MO9815657-A1.

XX 16-APR-1998.

XX 08-OCT-1997; 97WD-US018290.

XX 08-OCT-1996; 96US-00727688.

XX (ABBO ) ABBOTT LAB.

PI Cohen M, Friedman PN, Gordon J, Hodges SC, Klass MR;  
 PI Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD;  
 XX WPI; 1998-240838/21.

PT Detecting a target PS112 polynucleotide - used for diagnosing prostate  
 PT cancer.

XX Example 10; Page 92; 104pp; English.

CC AAW54426-W54433 represent synthetic peptide fragments of designed from  
 CC the human PS112 protein isolated from a prostate library. These peptides  
 CC are used in a novel method of detecting the presence of a target PS112  
 CC polynucleotide in a test sample. The method can also be used to detect  
 CC mRNA of PS112 in a test sample. The method can be used for diagnosis of  
 CC prostate cancer, as the presence of PS112 is an indicator of prostate  
 CC cancer. Antibodies against the polypeptides may be used as markers, or to  
 CC treat prostate cancer

XX Sequence 12 AA;

Query Match 33.3%; Score 18; DB 2; Length 12;  
 Best Local Similarity 28.6%; Pred. No. 6.3e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 DB 1 CSSSSSC 7

RESULT 14

AAB08422  
 ID AAB08422 standard; peptide; 12 AA.  
 XX  
 AC AAB08422;

DT 20-DEC-2000 (first entry)

XX Antigen peptide derived from prostate cancer associated protein PS112.  
 DE Prostate cancer associated gene; PS112; prostate disease;  
 XX prostate cancer; tumour; metastasis.  
 KW

XX Homo sapiens.  
 OS Synthetic.

XX US6110675-A.

XX 29-AUG-2000.

XX 08-OCT-1997; 97US-00946869.

XX 08-OCT-1996; 96US-00727688.

XX (ABBO ) ABBOTT LAB.

PI Friedman PN, Gordon J, Hodges SC, Klass MR, Cohen M,  
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H, Kratochvil JD;  
 XX WPI; 2000-571422/53.

XX 2000-571422/53.

PT Novel methods for diagnosing prostate cancer by contacting test sample  
 PT with target specific polynucleotide and detecting prostate cancer  
 PT associated polynucleotides.

XX Example 10; Col 77-78; 50pp; English.

CC AAB08416-23 represent antigenic peptides derived from a protein which is  
 CC encoded by a human prostate cancer associated gene, designated PS112. The  
 CC peptides are used to raise antibodies. PS112 sequences are useful for  
 CC detecting, diagnosing, staging, monitoring, prognosticating, preventing,  
 CC treating, or determining the predisposition of an individual to disease  
 CC and conditions of the prostate, such as prostate cancer, tumours and  
 CC metastases

XX Sequence 12 AA;

Query Match 33.3%; Score 18; DB 3; Length 12;  
 Best Local Similarity 28.6%; Pred. No. 6.3e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 DB 1 CSSSSSC 7

RESULT 15

ID ADQ96644 standard; peptide; 12 AA.

XX ADQ96644;

XX 23-SEP-2004 (first entry)

XX Human PS112 epitope seqid 32.

XX cytostatic; gene therapy; PS112; recombinant expression system;  
 KW PS112 epitope; prostate disease; tumours; metastasis; predisposition;  
 XX prostate cancer; epitope.

XX Synthetic.

XX US2004121397-A1.

Search completed: December 22, 2004, 06:30:36  
 Job time : 165 secs

PD 24-JUN-2004.  
 XX  
 PF 22-JAN-2004; 2004US-00763992.  
 XX  
 PR 08-OCT-1996; 96US-00727688.  
 PR 08-OCT-1997; 97US-00946869.  
 PR 15-OCT-1999; 99US-00418887.  
 XX  
 PA (COHE/) COHEN M.  
 PA (FRIE/) FRIEDMAN P N.  
 PA (GORD/) GORDON J.  
 PA (HODG/) HODGES S C.  
 PA (KIAS/) KIAS M R.  
 PA (KRAT/) KRATOCHVIL J D.  
 PA (ROBE/) ROBERTS-RAPP L.  
 PA (RUSS/) RUSSELL J C.  
 PA (STRO/) STROUPE S D.  
 PA (YUHH/) YU H.  
 XX  
 PI Cohen M, Friedman PN, Gordon J, Hodges SC, Kias MR,  
 PI Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;  
 XX  
 DR WPI: 2004-479676/45.  
 XX  
 PT Detecting a target PS112 polynucleotide, useful in diagnosing, staging,  
 PT monitoring, prognosticating, preventing and treating prostate cancer,  
 PT comprises contacting the test sample with PS112-specific polynucleotide.  
 XX  
 PS Example 10; SEQ ID NO 32; 53pp; English.  
 XX  
 CC The invention describes a method of detecting the presence of a target  
 CC PS112 polynucleotide in a test sample. The method comprises: contacting  
 CC the test sample with at least one PS112-specific polynucleotide or its  
 CC complement; and detecting the presence of the target PS112 polynucleotide  
 CC in the test sample, where the PS112-specific polynucleotide has at least  
 CC 50% identity to a polynucleotide comprising a sequence of 367, 214, 205,  
 CC 256, 246, 277, 251, 223, 2393, or 1297 bp (SEQ ID NOS: 1-10) or their  
 CC fragments or complements. Also described are: detecting mRNA of PS112 in  
 CC a test sample; a test kit, useful for detecting PS112 polynucleotide in  
 CC test sample; a purified polynucleotide or fragment derived from a PS112  
 CC gene; a recombinant expression system comprising a nucleic acid sequence  
 CC that includes an open reading frame derived from PS112 operably linked to  
 CC a control sequence compatible with a desired host, where the nucleic acid  
 CC sequence has at least 50% identity to a sequence of SEQ ID NOS: 1-10, or  
 CC their fragments or complements; a cell transfected with the recombinant  
 CC expression system or with a nucleic acid sequence encoding at least one  
 CC PS112 epitope, where the nucleic acid sequence comprises SEQ ID NOS: 1-  
 CC 10, or their fragments or complements; a composition of matter comprising  
 CC a PS112 polynucleotide or its fragment, where the polynucleotide has at  
 CC least 50% identity to a sequence of SEQ ID NOS: 2-10, or their  
 CC complements, or has at least 50% identity with fragments of a  
 CC polynucleotide of SEQ ID NOS: 4-8; and a gene or its fragment comprising  
 CC DNA having at least 50% identity with SEQ ID NOS: 9 or 10. The method is  
 CC useful for detecting the presence of a target PS112 polynucleotide in a  
 CC test sample. The methods, test kit, polynucleotides and polypeptides, and  
 CC antibodies are useful in detecting, diagnosing, staging, monitoring,  
 CC prognosticating, preventing and treating prostate diseases, tumours or  
 CC metastases or in determining the predisposition of an individual to  
 CC diseases and conditions of the prostate, e.g. prostate cancer. This is  
 CC the amino acid sequence of a PS112 epitope used in the creation of anti-  
 CC PS112-antibodies.  
 XX  
 XX Sequence 12 AA;  
 XX

Query Match: 33.3%; Score 18; DB 8; Length 12;

Best Local Similarity 28.6%; Pred. No. 6.3e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 | |  
 Db 1 CSSSSSC 7

**This Page Blank (uspto)**

Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 06:25:23 ; Search time 37 Seconds

(Without alignments)  
55.564 Million cell updates/sec

Title: US-09-732-348B-34

Perfect score: 54

Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXXXX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	33.3	15	4	US-09-142-524D-141
2	18	33.3	15	4	US-09-142-524D-142
3	18	33.3	24	4	US-09-731-242A-3
4	18	33.3	34	4	US-09-270-767-40588
5	18	33.3	34	4	US-09-270-767-55804
6	18	33.3	36	4	US-09-816-721-1
7	18	33.3	48	5	PCT-US96-01720-8
8	18	33.3	50	3	US-08-900-330-58
9	18	33.3	70	4	US-09-621-976-6084
10	18	33.3	90	4	US-09-419-381-90
11	18	33.3	125	4	US-09-252-991A-27726
12	18	33.3	127	3	US-08-467-023-189
13	18	33.3	132	4	US-09-252-991A-22556
14	18	33.3	134	4	US-09-252-991A-25069
15	18	33.3	138	4	US-09-252-991A-16583
16	18	33.3	139	4	US-09-252-991A-29513
17	18	33.3	141	4	US-09-825-994-215
18	18	33.3	141	4	US-09-970-966-215
19	18	33.3	143	4	US-09-252-991A-17322
20	18	33.3	148	4	US-09-252-991A-17476
21	18	33.3	148	4	US-09-252-991A-23009
22	18	33.3	152	4	US-09-252-991A-24277
23	18	33.3	156	4	US-09-252-991A-31875
24	18	33.3	158	4	US-09-252-991A-30633
25	18	33.3	163	2	US-08-727-688-25
26	18	33.3	163	4	US-09-252-991A-29098
27	18	33.3	166	4	US-09-252-991A-27896

#### ALIGNMENTS

28	18	33.3	174	4	US-09-252-991A-25719	Sequence 25719, A
29	18	33.3	178	4	US-09-252-991A-31386	Sequence 31386, A
30	18	33.3	190	4	US-09-252-991A-17994	Sequence 17994, A
31	18	33.3	193	4	US-09-252-991A-22723	Sequence 22723, A
32	18	33.3	195	4	US-09-270-767-36194	Sequence 36194, A
33	18	33.3	195	4	US-09-270-767-51411	Sequence 51411, A
34	18	33.3	249	4	US-09-252-991A-17158	Sequence 17158, A
35	18	33.3	278	4	US-09-252-991A-17001	Sequence 17001, A
36	18	33.3	320	4	US-09-252-991A-24634	Sequence 24634, A
37	18	33.3	341	2	US-08-209-521-11	Sequence 11, Appl
38	18	33.3	341	4	US-09-252-991A-20062	Sequence 20062, A
39	18	33.3	368	4	US-09-252-991A-26170	Sequence 26170, A
40	18	33.3	384	4	US-09-252-991A-24427	Sequence 24427, A
41	18	33.3	397	4	US-09-252-991A-18709	Sequence 18709, A
42	18	33.3	404	1	US-08-696-770-2	Sequence 2, Appl
43	18	33.3	404	2	US-09-015-557-2	Sequence 26942, A
44	18	33.3	437	4	US-09-252-991A-26942	Sequence 26942, A
45	18	33.3	441	4	US-09-328-352-6745	Sequence 6745, Ap

```

RESULT 1
US-09-142-524D-141
Sequence 141, Application US/09142524D
Patent No. 6719976
GENERAL INFORMATION:
APPLICANT: Some, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kuno, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 141
LENGTH: 15
TYPE: PRT
ORGANISM: Cryptomeria japonica
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1)-(15)
OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 58
US-09-142-524D-141

Query Match 33.3%; Score 18; DB 4; Length 15;
Best local similarity 28.6%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 8 CTSASAC 14

RESULT 2
US-09-142-524D-142
Sequence 142, Application US/09142524D
Patent No. 6719976
GENERAL INFORMATION:
APPLICANT: Some, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kuno, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D

```

```
; CURRENT FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: PCT/JP97/00740
; PRIOR FILING DATE: 1997-03-10
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 142
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Cryptomeria japonica
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 59
US-09-142-524D-142
```

```
Query Match          33.3%; Score 18; DB 4; Length 15;
Best Local Similarity 28.6%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 3 CXXXXXC 9
Db 3 CTSASAC 9
```

```
RESULT 3
US-09-731-242A-3
; Sequence 3, Application US/09731242A
; Patent No. 6759243
; GENERAL INFORMATION:
; APPLICANT: KRANZ, DAVID
; APPLICANT: WITTRUP, K. DANF
; APPLICANT: HOLTER, PHILIP
; TITLE OF INVENTION: HIGH AFFINITY TCR PROTEINS AND METHODS
; FILE REFERENCE: 89-99
; CURRENT APPLICATION NUMBER: US/09/731, 242A
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 60/169,179
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/009,388
; PRIOR FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 24
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: upstream primer
US-09-731-242A-3
```

```
Query Match          33.3%; Score 18; DB 4; Length 24;
Best Local Similarity 28.6%; Pred. No. 1.9e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 3 CXXXXXC 9
Db 9 CATTAAC 15
```

```
RESULT 4
US-09-270-767-40588
; Sequence 40588, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 40588
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-40588
```

```
Query Match          33.3%; Score 18; DB 4; Length 34;
Best Local Similarity 28.6%; Pred. No. 2.2e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 3 CXXXXXC 9
Db 4 CXXXXXC 10
```

```
RESULT 5
US-09-270-767-55804
; Sequence 55804, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55804
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-55804
```

```
Query Match          33.3%; Score 18; DB 4; Length 34;
Best Local Similarity 28.6%; Pred. No. 2.2e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 3 CXXXXXC 9
Db 4 CXXXXXC 10
```

```
RESULT 6
US-09-816-721-1
; Sequence 1, Application US/09816721
; Patent No. 6777215
; GENERAL INFORMATION:
; APPLICANT: No. 6777215ozymes A/S
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Haren, Lubbet D.
; APPLICANT: Dijkstra, Babke H.
; TITLE OF INVENTION: No. 6777215el Cyclomalodextrin glucanotransferase Variants
; FILE REFERENCE: 5347,210-US
; CURRENT APPLICATION NUMBER: US/09/816,721
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
; NAME/KEY: misc feature
; OTHER INFORMATION: "D196H oligo"
US-09-816-721-1
```

```
Query Match          33.3%; Score 18; DB 4; Length 36;
Best Local Similarity 28.6%; Pred. No. 2.3e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 3 CXXXXXC 9
```



Db 25 CTAATC 31

## RESULT 7

PCT-US96-01720-8

Sequence 8, Application PC/TUS9601720

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: MODIFIED-AFFINITY STREPTAVIDIN

NUMBER OF SEQUENCES: 11

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/01720

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/387,055

FILING DATE: 09-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 16336-5PC

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 48 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-01720-8

Query Match 33.3%; Score 18; DB 5; Length 48;

Best Local Similarity 28.6%; Pred. No. 2.7e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

Db 6 CTAATC 12

## RESULT 8

US-08-900-230-58

Sequence 58, Application US/08900230

Patent No. 6329197

GENERAL INFORMATION:

APPLICANT: Bard, Jonathan A.

TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND

TITLE OF INVENTION: USES THEREOF

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper &amp; Dunham LLP

STREET: 1185 Avenue of The Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 11036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/900,230

FILING DATE: 23-JUL-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE:

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-900-230-58

Query Match 33.3%; Score 18; DB 3; Length 50;

Best Local Similarity 28.6%; Pred. No. 2.7e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

Db 9 CTAATAC 15

## RESULT 9

US-09-621-976-6084

Sequence 6084, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 6084

LENGTH: 70

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: UNSURE

LOCATION: 68

OTHER INFORMATION: Xaa = Ala,Thr

NAME/KEY: UNSURE

LOCATION: 69

OTHER INFORMATION: Xaa = Gly,Arg

NAME/KEY: UNSURE

LOCATION: 18

OTHER INFORMATION: Xaa = Lys,Arg

US-09-621-976-6084

Query Match 33.3%; Score 18; DB 4; Length 70;

Best Local Similarity 28.6%; Pred. No. 3.3e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

Db 42 CSASTSC 48

## RESULT 10

US-09-419-381-90

Sequence 90, Application US/09419381

Patent No. 6747135

GENERAL INFORMATION:

APPLICANT: No. 6747135an, Garry P.

APPLICANT: Rozinov, Michael N.

TITLE OF INVENTION: Fluorescent Dye Binding Peptides

FILE REFERENCE: A65681-1/DJB/RMS/DSS

CURRENT APPLICATION NUMBER: US/09/419,381

```

; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/104,465
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-419-381-90

Query Match          33.3%; Score 18; DB 4; Length 90;
Best Local Similarity 28.6%; Pred. No. 3.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 45 CAAATC 51

RESULT 11
US-09-252-991A-27726
; Sequence 27726, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27726
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27726

Query Match          33.3%; Score 18; DB 4; Length 125;
Best Local Similarity 28.6%; Pred. No. 4.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 17 CTASTTC 23

RESULT 12
US-08-467-023-189
; Sequence 189, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griffech, Irwin J.;
; APPLICANT: Pollock, Joanne;
; APPLICANT: Bond, Julian F.;
; APPLICANT: Garman, Richard D.;
; APPLICANT: Kuo, Mei-Chang;
; APPLICANT: Yeung, Siu-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
```

```

; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 189:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-467-023-189

Query Match          33.3%; Score 18; DB 3; Length 127;
Best Local Similarity 28.6%; Pred. No. 4.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 77 CTASAC 83

RESULT 13
US-09-252-991A-22556
; Sequence 22556, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22556
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22556

Query Match          33.3%; Score 18; DB 4; Length 132;
Best Local Similarity 28.6%; Pred. No. 4.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 33 CAAASAC 39

RESULT 14
```

US-09-252-991A-25069  
; Sequence 25069; Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25069  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25069

Query Match 33.3%; Score 18; DB 4; Length 134;  
Best Local Similarity 28.6%; Pred. No. 4.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 11 CSASTTC 17

RESULT 15  
US-09-252-991A-16583  
; Sequence 16583; Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16583  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16583

Query Match 33.3%; Score 18; DB 4; Length 138;  
Best Local Similarity 28.6%; Pred. No. 4.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 53 CASTTAC 59

Search completed: December 22, 2004, 06:35:21  
Job time: 40 secs

This Page Blank (usp<sup>ts</sup>)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 06:27:53 ; Search time 143 Seconds

(without alignments)  
77.573 Million cell updates/sec

Title: US-09-732-348B-34

Perfect score: 54  
Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXXXX 31

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	35.2	103	10	US-09-820-843A-77
2	18	33.3	10	10	US-09-572-404B-2385
3	18	33.3	10	10	US-09-572-404B-2387
4	18	33.3	10	10	US-09-572-404B-2399
5	18	33.3	10	10	US-09-572-404B-2401
6	18	33.3	10	10	US-09-572-404B-2425
7	18	33.3	10	10	US-09-572-404B-2485
8	18	33.3	10	10	US-09-572-404B-2527
9	18	33.3	10	10	US-09-572-404B-2533
10	18	33.3	10	10	US-09-572-404B-2543
11	18	33.3	10	10	US-09-572-404B-2557
12	18	33.3	12	16	US-10-763-992-32
13	18	33.3	15	14	US-10-354-240-141

14	18	33.3	15	14	US-10-354-240-142	Sequence 142, App
15	18	33.3	18	14	US-10-225-567A-1778	Sequence 1778, App
16	18	33.3	24	9	US-09-731-242A-3	Sequence 3, Appl
17	18	33.3	25	14	US-10-361-811-292	Sequence 292, App
18	18	33.3	25	14	US-10-369-186-292	Sequence 292, App
19	18	33.3	28	9	US-09-864-761-41104	Sequence 41104, A
20	18	33.3	33	16	US-10-697-399-10	Sequence 10, Appl
21	18	33.3	36	9	US-09-816-721-1	Sequence 1, Appl
22	18	33.3	39	14	US-10-029-386-31509	Sequence 31509, A
23	18	33.3	40	17	US-10-660-206-83	Sequence 83, Appl
24	18	33.3	50	8	US-08-900-230-58	Sequence 58, Appl
25	18	33.3	53	17	US-10-425-115-321769	Sequence 321769, Sequence 353488,
26	18	33.3	53	17	US-10-425-115-353488	Sequence 205445,
27	18	33.3	57	17	US-10-425-115-205445	Sequence 28972, A
28	18	33.3	58	14	US-10-029-386-28972	Sequence 276806,
29	18	33.3	63	17	US-10-425-115-276806	Sequence 193586,
30	18	33.3	66	16	US-10-437-963-193586	Sequence 202235,
31	18	33.3	67	15	US-10-424-599-202235	Sequence 196666,
32	18	33.3	67	17	US-10-425-115-196666	Sequence 261219,
33	18	33.3	70	15	US-10-424-599-261219	Sequence 32, Appl
34	18	33.3	71	14	US-10-011-931-32	Sequence 45, Appl
35	18	33.3	71	14	US-10-269-806-45	Sequence 61, Appl
36	18	33.3	71	14	US-10-269-806-61	Sequence 65, Appl
37	18	33.3	71	14	US-10-269-806-65	Sequence 77, Appl
38	18	33.3	71	14	US-10-269-806-77	Sequence 85, Appl
39	18	33.3	71	14	US-10-269-806-85	Sequence 109, App
40	18	33.3	71	14	US-10-269-806-109	Sequence 117, App
41	18	33.3	71	14	US-10-269-806-117	Sequence 133, App
42	18	33.3	71	14	US-10-269-806-133	Sequence 11, Appl
43	18	33.3	75	9	US-09-758-140-11	Sequence 199736,
44	18	33.3	77	16	US-10-437-963-199736	Sequence 107, App
45	18	33.3	81	15	US-10-131-487A-107	

#### ALIGNMENTS

RESULT 1  
US-09-820-843A-77  
; Sequence 77, Application US/09820843A  
; Publication No. US2003003963A1  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 77  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: M. pneumoniae  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: B01\_orf103b Protein  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: gi|1673772  
US-09-820-843A-77

Query Match 35.2%; Score 19; DB 10; Length 103;  
Best Local Similarity 13.6%; Pred. No. 5.5e+03;  
Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 3 CXXXXXXXXXXXXXXXXXXXXH 24  
DB 53 CCSIIFCSLASSARLRYSSSH 74

RESULT 2  
US-09-572-404B-2385  
; Sequence 2385, Application US/09572404B

```

; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2385
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 82-91 and may interact with Sequence 2
US-09-572-404B-2385

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 3 CSSSTTC 9

RESULT 3
US-09-572-404B-2387
; Sequence 2387, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2387
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-09-572-404B-2387

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 4
US-09-572-404B-2399
; Sequence 2399, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2399
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
```

```

; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-09-572-404B-2399

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 5
US-09-572-404B-2401
; Sequence 2401, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2401
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-09-572-404B-2401

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 2 CSSSTTC 8

RESULT 6
US-09-572-404B-2425
; Sequence 2425, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2425
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 84-93 and may interact with Sequence 2
US-09-572-404B-2425

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 1 CSSSTTC 7
```

```
RESULT 7
US-09-572-404B-2485
; Sequence 2485, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2485
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-09-572-404B-2485

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Pred. No. 2.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 4 CSSSTTC 10

RESULT 8
US-09-572-404B-2527
; Sequence 2527, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2527
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-09-572-404B-2527

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Pred. No. 2.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 2 CSSSTTC 8

RESULT 9
US-09-572-404B-2533
; Sequence 2533, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
```

```
; SEQ ID NO 2533
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-09-572-404B-2533

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Pred. No. 2.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 2 CSSSTTC 8

RESULT 10
US-09-572-404B-2543
; Sequence 2543, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2543
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-09-572-404B-2543

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Pred. No. 2.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 4 CSSSTTC 10

RESULT 11
US-09-572-404B-2557
; Sequence 2557, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2557
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-09-572-404B-2557

Query Match
Best Local Similarity 33.3%; Score 18; DB 10; Length 10;
Pred. No. 2.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

QY 3 CXXXXXC 9  
DB 2 CXXXXXC 8

RESULT 12  
US-10-763-992-32  
Sequence 32, Application US/10763992  
Publication No US20040121397A1  
GENERAL INFORMATION:

APPLICANT: COHEN, Maurice  
FRIEDMAN, Paula N.  
GORDON, Julian  
HODGES, Steven C.  
KLASS, Michael R.  
KRATOCHVIL, Jon D.  
ROBERTS-RAPP, Lisa  
RUSSELL, John C.  
STROUPE, Steven D.  
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL  
FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL

COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/763,992  
FILING DATE: 22-Jan-2004  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/418,887  
FILING DATE: 15-OCT-1999  
APPLICATION NUMBER: US/08/946,869  
FILING DATE: 08-OCT-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 5697.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: None  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-10-763-992-32

Query Match 33.3%; Score 18; DB 16; Length 12;  
Best Local Similarity 28.6%; Pred. No. 3e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 1 CXXXXXC 7

RESULT 13  
US-10-354-240-141  
Sequence 141, Application US/10354240

Publication No. US20030185847A1

GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103D1

CURRENT APPLICATION NUMBER: US/10/354,240

PRIOR FILING DATE: 1997-03-10

PRIOR APPLICATION NUMBER: US 09/142,524

PRIOR FILING DATE: 1998-09-09

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 141

LENGTH: 15

TYPE: PRT

ORGANISM: Cryptomeria japonica

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (1)-(15)

OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 58

Query Match 33.3%; Score 18; DB 14; Length 15;  
Best Local Similarity 28.6%; Pred. No. 3.3e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 8 CTSASAC 14

RESULT 14

US-10-354-240-142

Sequence 142, Application US/10354240

Publication No. US20030185847A1

GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103D1

CURRENT APPLICATION NUMBER: US/10/354,240

PRIOR FILING DATE: 2003-01-29

PRIOR APPLICATION NUMBER: PCT/J997/00740

PRIOR FILING DATE: 1997-03-10

PRIOR APPLICATION NUMBER: US 09/142,524

PRIOR FILING DATE: 1998-09-09

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 142

LENGTH: 15

TYPE: PRT

ORGANISM: Cryptomeria japonica

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (1)-(15)

OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 59

Query Match 33.3%; Score 18; DB 14; Length 15;  
Best Local Similarity 28.6%; Pred. No. 3.3e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 3 CTSASAC 9



## RESULT 15

US-10-225-567A-1778

; Sequence 1778, Application US/10225567A

; Publication No. US20030113798A1

## GENERAL INFORMATION:

; APPLICANT: LifeSpan Biosciences

; APPLICANT: Brown, Joseph P.

; APPLICANT: Burmet, Glenna C.

; APPLICANT: Roush, Christine L.

; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 1920-4-4

; CURRENT APPLICATION NUMBER: US/10/225,567A

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/257,144

; PRIOR FILING DATE: 2000-12-19

; NUMBER OF SEQ ID NOS: 2292

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO: 1778

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-225-567A-1778

Query Match 33.3%; Score 18; DB 14; Length 18;

Best Local Similarity 28.6%; Pred. No. 3.7e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

DB 9 CASSSC 15

Search completed: December 22, 2004, 06:37:49  
Job time : 145 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 06:24:38 / Search time 39 Seconds

(Without alignments)  
76,480 Million cell updates/sec

Title: US-09-732-348B-34

Perfect score: 54  
Sequence: 1 XXCXXXXXXCXXXXXXCXXXXXXCXXXXXX 31

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR 79: \*  
2: PIR1: \*  
3: PIR2: \*  
4: PIR3: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35.2	103	2	S73440	hypothetical prote
2	33.3	121	1	RMBEM3	surface glycoprote
3	33.3	135	2	A45835	Ly6 homolog RK10 p
4	33.3	157	2	S09805	hypothetical prote
5	33.3	295	2	T00399	probable AP2 domai
6	33.3	381	2	B97542	hypothetical prote
7	33.3	381	2	AD2761	rare lipoprotein A
8	33.3	404	2	UC5784	adrenomedullin rec
9	33.3	442	2	T24196	hypothetical prote
10	33.3	492	2	S49147	RRF-2 protein - hu
11	33.3	507	2	UC7366	Jun A 2 protein -
12	33.3	514	2	S48730	Cty j II protein -
13	33.3	514	2	UC2498	second major aller
14	33.3	514	2	UC7100	polygalacturonase
15	33.3	525	2	IS3194	cdc25A - rat
16	33.3	584	1	C8HUA	complement C8 alph
17	33.3	675	2	D48492	kpsc protein - Bac
18	33.3	746	2	T24578	hypothetical prote
19	33.3	1817	2	T34249	hypothetical prote
20	33.3	25	1	SMNR	metallochionein -
21	31.5	26	1	SMNC	metallochionein -
22	31.5	26	2	SS5029	CAP3 protein - ant
23	31.5	27	2	S42359	ferredoxin [3Fe-4S
24	31.5	45	1	PFY2	fluvocin C - Myxoc
25	31.5	45	2	AC3539	hypothetical prote
26	31.5	61	1	SMHY1C	metallochionein I
27	31.5	61	2	S00810	metallochionein I
28	31.5	66	2	SS9621	metallochionein is
29	31.5	71	2	H95299	hypothetical prote

30	17	31.5	76	2	D44007	apoptoxin IV - tra
31	17	31.5	76	2	C44007	apoptoxin VI - tra
32	17	31.5	78	2	E81853	probable transposa
33	17	31.5	89	2	S55780	Hox C6 protein - e
34	17	31.5	93	2	E83307	hypothetical prote
35	17	31.5	93	2	A97190	hypothetical prote
36	17	31.5	100	2	T17962	hypothetical prote
37	17	31.5	103	2	G84741	hypothetical prote
38	17	31.5	108	2	S17201	protein kinase (EC
39	17	31.5	110	2	D72701	hypothetical prote
40	17	31.5	117	2	UC2210	hypothetical 12.6
41	17	31.5	128	1	A57321	E48 antigen precu
42	17	31.5	128	2	G97176	stress-induced pro
43	17	31.5	130	2	A54762	phospholipase A2 (
44	17	31.5	131	2	A43980	neurophysin 2 (val
45	17	31.5	131	2	I56894	complement regulat

#### ALIGNMENTS

RESULT 1  
S73440  
hypothetical protein B01.orf103b - Mycoplasma pneumoniae (strain ATCC 29342)  
A:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C>Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S73440  
R:Himmelreich, R.; Hilbert, H.; Piagans, H.; Pirkli, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A/Reference number: S73327; MUID:97105885; PMID:8948633  
A/Accession: S73440  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-103 <HM>  
A/Cross-references: UNIPROT:P75074; EMBL:AE000014; GB:U00089; NID:g1673770; PIDN:AA9576  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C/Genetics:  
A/Genetic code: SGC3

Query Match 35.2% Score 19; DB 2; Length 103;  
Best Local Similarity 13.6% Pred. No. 2.4e+02;  
Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 3 CXXXXXXCXXXXXXCXXXXXXH 24  
DB 53 CSSISFCSLASSSARLRYSSTH 74

RESULT 2  
RMBEM3  
surface glycoprotein CD59 precursor homolog - simlirine herpesvirus 1 (strain 11)  
C/Species: simlirine herpesvirus 1  
A/Note: host Simlirine scuturus (common squirrel monkey)  
C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 22-Jun-1999  
C/Accession: A43384; F36807  
R:Albrecht, J.C.; Nicholas, J.; Cameron, K.R.; Newman, C.; Fleckenstein, B.; Honess, R.W  
Virology 190, 527-530, 1992  
A>Title: Herpesvirus simliri has a gene specifying a homologue of the cellular membrane  
A/Reference number: A43384; MUID:92410640; PMID:1382344  
A/Accession: A43384  
A/Molecule type: DNA  
A/Residues: 1-121 <ALB>  
A/Cross-references: GB:S44811; GB:X64273; NID:g60318; PIDN:CAA45565.1; PID:g60319  
C/Genetics:  
A/Genes: 15  
C/Superfamily: Ly-6 antigen; Ly-6 homology  
C/Keywords: glycoprotein; lipoprotein; phosphatidylinositol linkage  
F:1-19/Domain: signal sequence #status predicted <Sig>  
F:20-96/Product: surface glycoprotein #status predicted <SGP>  
F:97-121/Domain: carboxyl-terminal propeptide #status predicted <CPP>

F:24/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.3%; Score 18; DB 1; Length 121;  
Best Local Similarity 28.6%; Pred. No. 4.8e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 32 CTTSTSC 38

## RESULT 3

A45835

Ly6 homolog RK10 precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 03-Mar-1994 #sequence \_revision 03-Mar-1994 #text\_change 09-Jul-2004

C:Accession: A45835; B45835

R:Friedman, S.; Palfrey, R.G.E.; Stirling, S.; Haemmerling, U.

Immunogenetics 31, 104-111, 1990

A:Title: Analysis of three distinct Ly6-A-related cDNA sequences isolated from rat kidney

A:Reference number: A45835; PMID:2154400

A:Accession: A45835

A:Molecule type: mRNA

A:Residues: 1-135 &lt;FRI&gt;

A:Cross-references: UNIPROT:O63317; GB:M30689; NID:G205247; PIDN:AAA1546.1; PID:G205248

A:Experimental source: clone RK10

A:Accession: B45835

A:Molecule type: mRNA

A:Residues: 25-54, 'A', 56-62, 'W', 64-67, 'Q', 69-71, 'DHI', 75-77, 'V', 79-80, 'T', 82-85, 'T', 87, 'A:Cross-references: GB:M30692; NID:G205245; PIDN:AAA1545.1; PID:G205246

A:Experimental source: clone RK6

C:Superfamily: Ly-6 antigen; Ly-6 homolog

C:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linkage

F:1-6/Domain: signal sequence #status predicted &lt;Strg&gt;

F:106/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asn) (in mature form)

Query Match 33.3%; Score 18; DB 2; Length 135;  
Best Local Similarity 28.6%; Pred. No. 5e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 41 CSTRATC 47

## RESULT 4

S09805

hypothetical protein UL42 - human cytomegalovirus (strain AD169)

C:Species: human cytomegalovirus, human herpesvirus 5

C:Date: 07-Sep-1990 #sequence \_revision 07-Sep-1990 #text\_change 09-Jul-2004

C:Accession: S09805

R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornsneil, T.;

M.; Barrell, B.G. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A:Reference number: S09749; PMID:90269039; PMID:2161319

A:Accession: S09805

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-157 &lt;CHE&gt;

A:Cross-references: UNIPROT:P16815; EMBL:X17403; NID:G59591; PIDN:CAA35401.1; PID:e27241

A:Note: this sequence was submitted to the EMBL Data Library, December 1989

A:Note: this reading frame extends between two stop codons and does not begin with a start

C:Superfamily: human cytomegalovirus hypothetical protein UL42

C:Keywords: glycoprotein; transmembrane protein

F:117-143/Domain: transmembrane #status predicted &lt;TM&gt;

F:147/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.3%; Score 18; DB 2; Length 157;  
Best Local Similarity 28.6%; Pred. No. 5.4e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 76 CAATSSC 82

## RESULT 5

T00399

probable AP2 domain transcription factor [imported] - Arabidopsis thaliana

N:Alternate names: transcription factor TINY homolog T13E15.5

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 01-Feb-1999 #sequence \_revision 01-Feb-1999 #text\_change 09-Jul-2004

C:Accession: T00399; F84884

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Masor

submitted to the EMBL Data Library, July 1997

A:Description: Arabidopsis thaliana chromosome II BAC T13E15 genomic sequence.

A:Reference number: Z14146

A:Accession: T00399

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-295 &lt;RCU&gt;

A:Cross-references: UNIPROT:O22158; EMBL:AC002388; NID:G3420042; PID:G3444890

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; PMID:20083487; PMID:10617197

A:Accession: F84884

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-295 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:G4895256; PIDN:AAD32841.1; GSPDB:GN00139

C:Genetics:

A:Gene: T13E15.5; AT2G44940

A:Map position: 2

Query Match 33.3%; Score 18; DB 2; Length 295;  
Best Local Similarity 28.6%; Pred. No. 7.1e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 233 CASTSAC 239

## RESULT 6

B97542

hypothetical protein AGR C.2765 [imported] - Agrobacterium tumefaciens (strain C58, Cerec

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence \_revision 30-Sep-2001 #text\_change 09-Jul-2004

C:Accession: B97542

R:Goodier, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97353; PMID:21608551; PMID:11743194

A:Accession: B97542

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-381 &lt;KUR&gt;

A:Cross-references: UNIPROT:Q8UP98; GB:AE007869; PIDN:AAK87291.1; PID:G15156585; GSPDB:GI

C:Genetics:

A:Gene: AGR C.2765

A:Map position: circular chromosome

C:Superfamily: p1A lipoprotein

Query Match 33.3%; Score 18; DB 2; Length 381;  
Best Local Similarity 28.6%; Pred. No. 8e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

Db 44 CAATASC 50

# RESULT 7

AD2761 rare lipoprotein A [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C/Species: Agrobacterium tumefaciens

C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C/Accession: AD2761

C/Mod: D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A/Reference number: AB2577; MUID:21608550; PMID:11743193

A/Accession: AD2761

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-381 <KOR>

A/Cross-references: UNIPROT:Q8UF98; GB:AE008688; PIDN:AAU42506.1; PID:917739925; GSPDB:C

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Map position: circular chromosome

C/Superfamily: rplA lipoprotein

Query Match 33.3%; Score 18; DB 2; Length 381;  
Best Local Similarity 28.6%; Pred. No. 8e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
Db 44 CAATASC 50

# RESULT 8

JC5784 adrenomedullin receptor - human

C/Species: Homo sapiens (man)

C/Date: 24-Jan-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004

C/Accession: JC5784

R/haene, J.; Dietrich, K.; Doetsch, J.; Raecher, W.

Biochem. Biophys. Res. Commun. 240, 183-188, 1997

A/Title: Molecular cloning of a novel human receptor gene with homology to the rat adren

A/Reference number: JC5784; MUID:98042541; PMID:9367907

A/Accession: JC5784

A/Molecule type: mRNA

A/Residues: 1-404 <HAE>

A/Cross-references: UNIPROT:O15218; GB:Y13583; NID:G2652933; PIDN:CAA73910.1; PID:G26529

C/Superfamily: vertebrate rhodopsin

C/Keywords: glycoprotein; receptor; transmembrane protein

F/58-79/Domain: transmembrane #status predicted <TM1>

F/91-113/Domain: transmembrane #status predicted <TM2>

F/128-149/Domain: transmembrane #status predicted <TM3>

F/171-193/Domain: transmembrane #status predicted <TM4>

F/218-239/Domain: transmembrane #status predicted <TM5>

F/260-280/Domain: transmembrane #status predicted <TM6>

F/301-320/Domain: transmembrane #status predicted <TM7>

F/28,37/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 33.3%; Score 18; DB 2; Length 404;  
Best Local Similarity 28.6%; Pred. No. 8.2e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
Db 349 CASSASC 355

# RESULT 9

T24196

hypothetical protein R11H6.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T24196

R/Bardill, S.

submitted to the EMBL Data Library, March 1997

A/Reference number: Z19852

A/Accession: T24196

A/Status: preliminary; translated from GB/EMBL/DBD

A/Molecule type: DNA

A/Residues: 1-442 <MTL>

A/Cross-references: UNIPROT:Q45719; EMBL:Z93386; PIDN:CAE07645.1; GSPDB:GNO0023; CESP:R11

A/Experimental source: clone R11H6

C/Genetics:

A/Map position: 5

A/Introns: 13/3; 68/3; 154/1; 284/1; 356/1

Query Match 33.3%; Score 18; DB 2; Length 442;  
Best Local Similarity 28.6%; Pred. No. 8.6e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
Db 10 CAASAC 16

# RESULT 10

S49147 ERF-2 protein - human

C/Species: Homo sapiens (man)

C/Date: 16-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004

C/Accession: S49147; I37372

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.

submitted to the EMBL Data Library, April 1994

A/Reference number: S49147

A/Accession: S49147

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-492 <RES>

A/Cross-references: UNIPROT:P47974

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.

Gene 152, 285-286, 1995

A/Title: ERF-2, the human homologue of the murine Tis11d early response gene.

A/Reference number: I37372; MUID:95137407; PMID:7835719

A/Accession: I37372

A/Status: preliminary; translated from GB/EMBL/DBD

A/Molecule type: mRNA

A/Residues: 1-492 <RES>

A/Cross-references: EMBL:X78992; NID:G509777; PIDN:CAA55592.1; PID:G509778

C/Genetics:

A/Map position: 13q12.3-13q12.3

Query Match 33.3%; Score 18; DB 2; Length 492;  
Best Local Similarity 28.6%; Pred. No. 9e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
Db 293 CASSASC 299

# RESULT 11

JC7366

Jun a 2 protein - mountain cedar

C/Species: Juniperus ashei (mountain cedar)

C/Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 09-Jul-2004

C/Accession: JC7366; PC7093

R/Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major alle

```

A:Reference number: JC7366
A:Molecule type: mRNA
A:Residues: 1-507 <YOK>
A:Cross-references: UNIPROT:Q9FYI9, GB:AJ404653
A:Accession: PC7093
A:Molecule type: protein
A:Residues: 55-63 <YOZ>
C:Comment: This protein, a second major allergen of mountain cedar pollen, which is involved in the polygalacturonase family.
C:Keywords: glycoprotein, pollen

Query Match      33.3%; Score 18; DB 2; Length 507;
Best Local Similarity    28.6%; Pred. No. 9.1e+02;
Matches     2; Conservative   0; Mismatches    5; Indels    0; Gaps    0;

OY       3 CXXXXXXC 9
Db        348 CTSAAAC 354

RESULT 12
S48730
CRY j II protein - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C>Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C:Accession: S48730
R:Namba, M., Kurose, M.; Tortigoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kurita, T.
FEBS Lett. 353, 124-128, 1994
A>Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar Picea nana.
A:Reference number: S48730, MUID:95010777, PMID:7926035
A:Accession: S48730
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-514 <NMA>
A:Cross-references: UNIPROT:P43212; GB:D37765; NID:g577695; PID:BAA07021.1; PID:d1007595

Query Match      33.3%; Score 18; DB 2; Length 514;
Best Local Similarity    28.6%; Pred. No. 9.2e+02;
Matches     2; Conservative   0; Mismatches    5; Indels    0; Gaps    0;

OY       3 CXXXXXXC 9
Db        347 CTSAAC 353

RESULT 13
JC2498
second major allergen Cry j II precursor - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C>Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: JC2498; PC2346; A60147
R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem Biophys Res Commun. 201, 1021-1028, 1994
A>Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese cedar.
A:Reference number: JC2498, MUID:94271186, PMID:8002972
A:Accession: JC2498
A:Molecule type: mRNA
A:Residues: 1-514 <KOM>
A:Cross-references: UNIPROT:P43212; DDBJ:D29772; NID:g506857; PIDN:BAA06172.1; PID:g506857
A:Accession: PC2346
A:Molecule type: protein
A:Residues: 52-61 <KOZ>
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhashi, T.
Allergy 45, 309-312, 1990
A>Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147, MUID:9034288; PMID:2382797
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAK>
C:KeyWords: glycoprotein, pollen
P:1-54/Domain: signal sequence #status predicted <SIG>
F:55-460/Product: second major allergen Cry j II #status predicted <MAT>
```

```

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          33.3%;   Score 18;   DB 2;   Length 514;
Best Local Similarity 28.6%;   Pred. No. 9,2e+02;
Matches 2;   Conservative 0;   Mismatches 5;   Indels 0;   Gaps 0;

QY          3 CXXXXXC 9
|
DB          347 CTSASAC 353

RESULT 14
JC7100
polygalacturonase Cha o 2 - Japanese cypress
C|Species: Chamaecyparis obtusa (Japanese cypress)
C|Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C|Accession: JC7100; PC7026
R|Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
B|Biochem. Biophys. Res. Commun. 263, 166-171, 1999
A|Title: Purification, identification, and cDNA cloning of Cha o 2, the second major allel
A|Reference number: JC7100; MUID:99417540; PMID:10486272
A|Accession: JC7100
A|Molecule type: mRNA
A|Residues: 1-514 <MOR>
A|Cross-references: UNIPROT:Q7M1E7
A|Accession: PC7026
A|Molecule type: protein
A|Residues: 51-62 <MO2>

Query Match          33.3%;   Score 18;   DB 2;   Length 514;
Best Local Similarity 28.6%;   Pred. No. 9,2e+02;
Matches 2;   Conservative 0;   Mismatches 5;   Indels 0;   Gaps 0;

QY          3 CXXXXXC 9
|
DB          347 CTSASAC 353

RESULT 15
153194
cdc25A - rat
C|Species: Rattus sp. (rat)
C|Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 22-Jun-1999
C|Accession: 153194
R|Jinno, S.; Suto, K.; Nagata, A.; Igarashi, M.; Kanoka, Y.; Nojima, H.; Okayama, H.
B|EMBO J. 13, 1549-1556, 1994
A|Title: Cdc25A is a novel phosphatase functioning early in the cell cycle.
A|Reference number: 153194; MUID:94208523; PMID:8156993
A|Accession: 153194
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: mRNA
A|Residues: 1-525 <RES>
A|Cross-references: GB:D16236; NID:9484291; PIDD:BA03761.1; PID:91008041
C|Superfamily: human protein-tyrosine-phosphatase cdc25A; cdc25-type protein-tyrosine-phosphatase homology <PTP>
F:314-507/Domain: cdc25-type protein-tyrosine-phosphatase homology <PTP>

Query Match          33.3%;   Score 18;   DB 2;   Length 525;
Best Local Similarity 28.6%;   Pred. No. 9,2e+02;
Matches 2;   Conservative 0;   Mismatches 5;   Indels 0;   Gaps 0;

QY          3 CXXXXXC 9
|
DB          266 CSSTSSC 272

Search completed: December 22, 2004, 06:34:37
Job time : 41 secs

```